Data-to-Decisions

Dr. Randy K. Avent

Data-to-Decision Systems

Tactical Operations

- Low Latency
- Narrow Field-of-View
- Limited Fusion
- Automatic Target Recognition
- Data: ~MB-GB

Operations Intelligence

- Medium Latency
- Wide Field-of-View
- Hard Sensor Fusion
- Assisted Target Recognition
- Data: ~GB-TB

Strategic Intelligence

- Long Latency
- Synoptic Field-of-View
- Hard/Soft Sensor Fusion
- Multiple Hypotheses
- Data: ~PB-EB

The complexity and adaptability of threats has surpassed our ability to find them in large data volumes within mission timelines

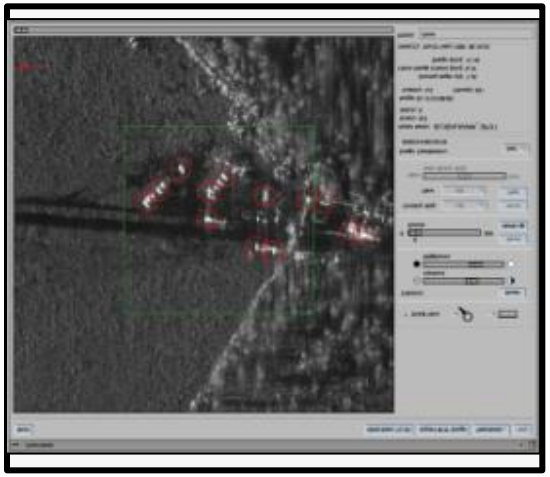
D2D Technology Assessment

- **Moderately Mature**
- **Driven by IT Industry**

Data Management Layer

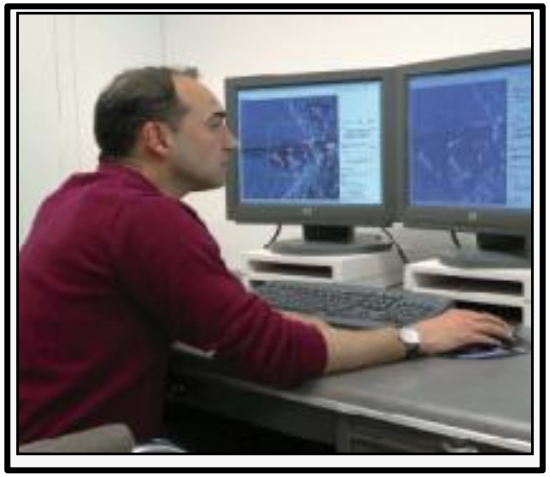
- **Immature**
- **Driven by Defense**

Analytics Layer



- **Moderately Mature**
- **Driven by IT Industry**

User Interface Layer



Current assessment is that unstructured data analytics is the most challenging and critical component of D2D

Outline

- Introduction
- **Technology Thrusts**
- **Summary**

Data Management Layer

- **Problem Statement: Increasing data volumes and modalities have diminished our ability to communicate, store, retrieve and process sources within mission critical timelines**

- **3-to-5 year timeframe objective**
  - Computational infrastructure to support capturing, processing, marking, retrieval, and management of millions of information objects per second
  - Network architecture with embedded information management on existing networks to support both real-time and discovery mission data requirements

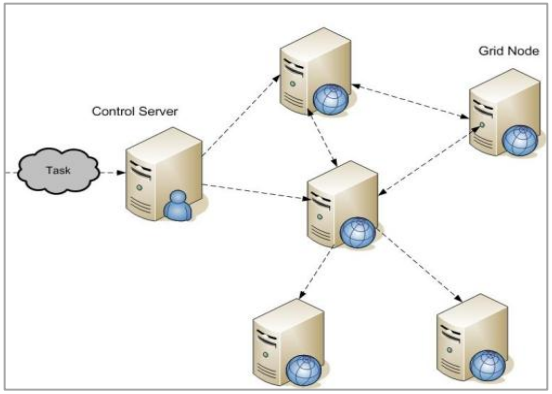
- **7-to-10 year timeframe objective**
  - Anticipatory autonomous control of sensors and compute resources to simultaneously support hundreds of consumer requests for analysis products

Hardware Infrastructure

• Embedded System

- On-board storage
- Tightly coupled data and algorithms
- Low-latency, low-bandwidth operations

• Grid Cluster



- Centralized storage
- Data moved to compute nodes
- Tightly coupled algorithms
- Parallel file system limits large data use

• Cloud Computing

- Distributed storage
- Applications moved to compute nodes
- Order-independence through map/reduce

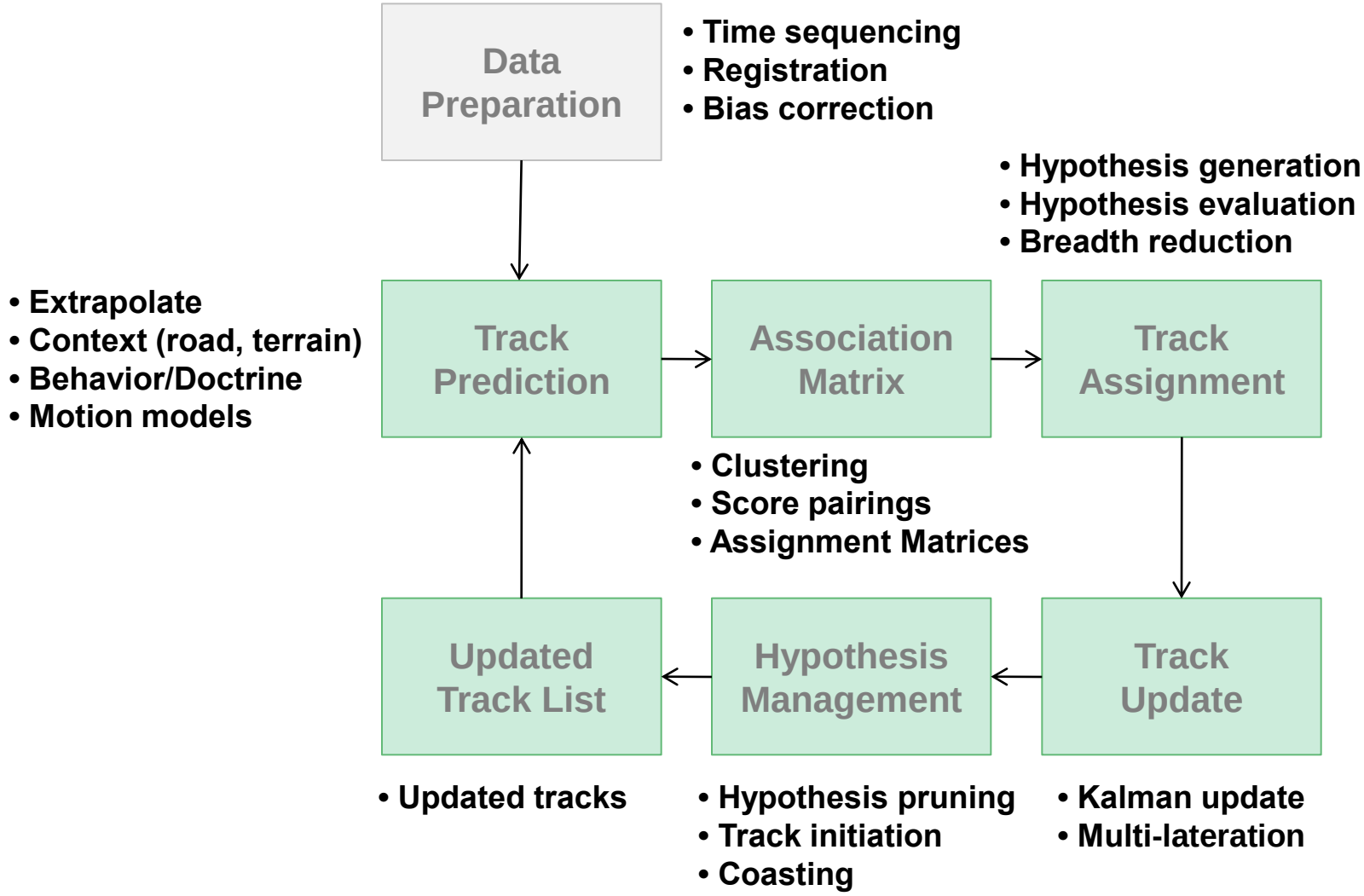
Analytic Layer

- **Problem Statement: Existing automation tools do not aid users in finding today's complex and adaptable threats within mission timelines**

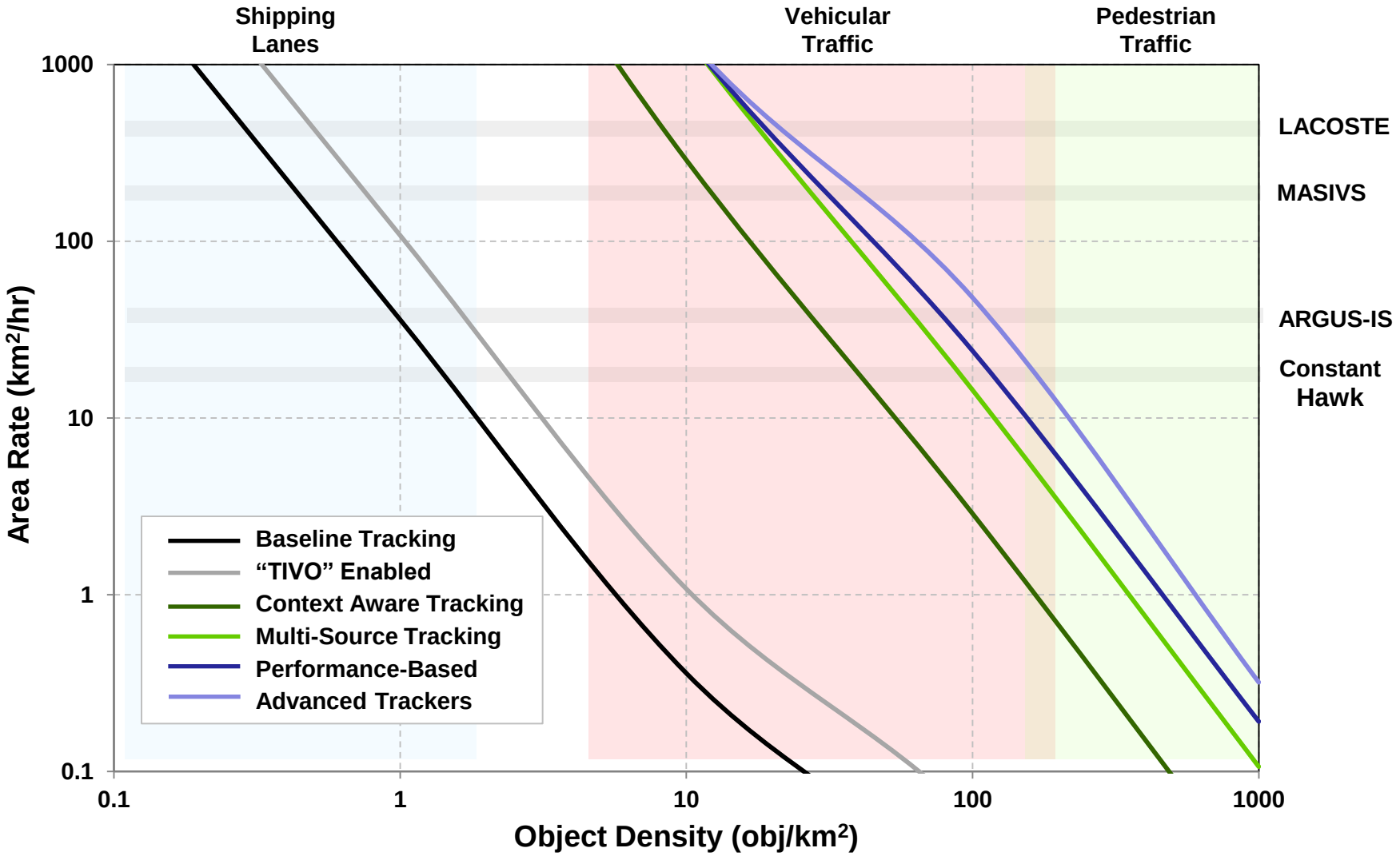
- **3-to-5 year timeframe objective**
  - Robust classification to accurately detect, geo-register and identify surface objects despite difficult environments, configurations and emplacements
  - Robust automation tools to identify relationships, patterns of life and activities of ground vehicles
  - Robust tools to capture, store and retrieve HUMINT-based information to identify and leverage popular support against insurgents

- **7-to-10 year timeframe objective**
  - Robust classification to accurately detect, geo-register and identify all surface objects despite difficult environments, configurations and emplacements
  - Robust automation tools to identify relationships, patterns of life and activities of dismounts
  - Robust tools to search, mine and exploit open-source data to identify all aspects of insurgent networks

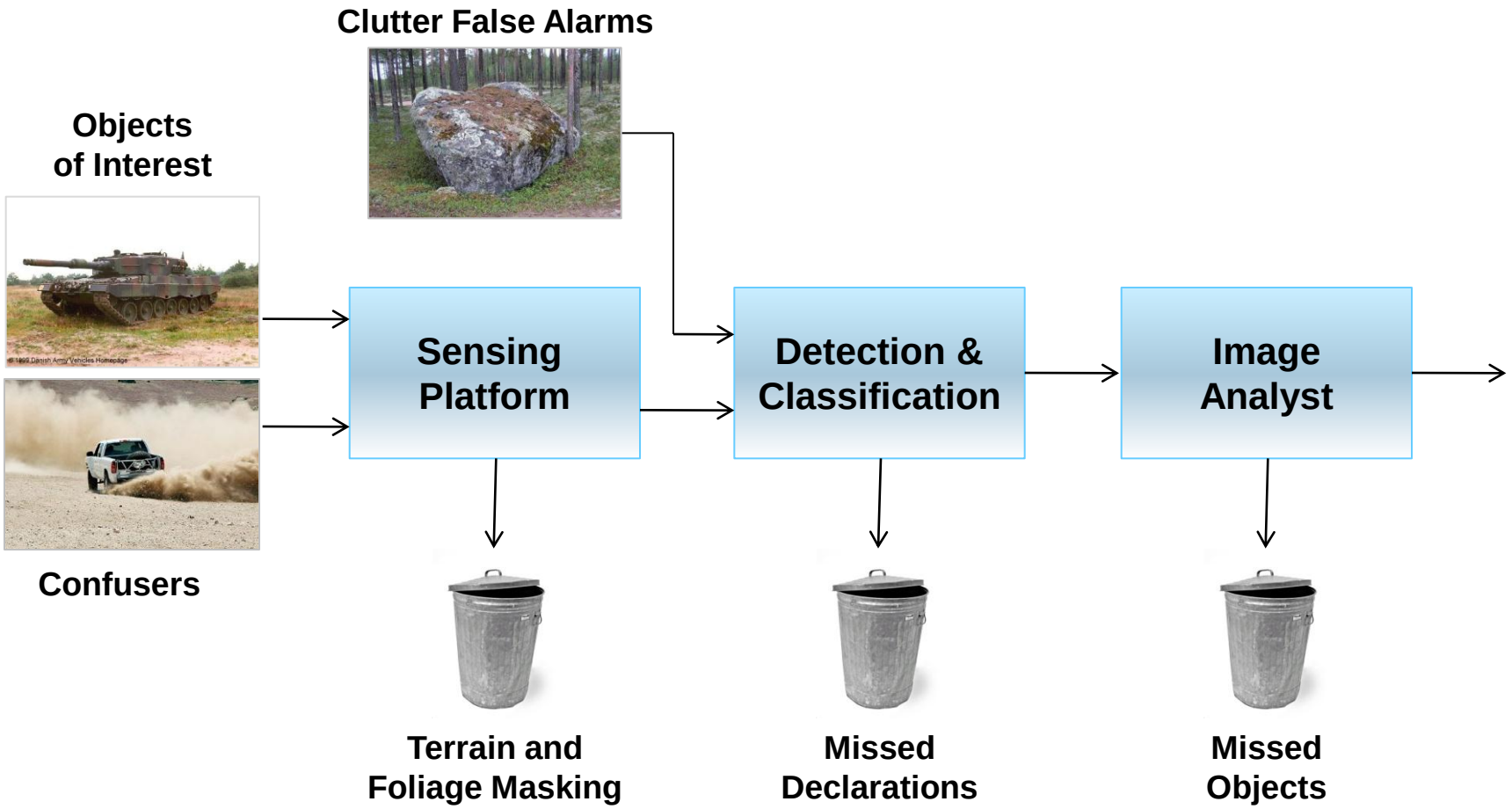
Generalized Tracking



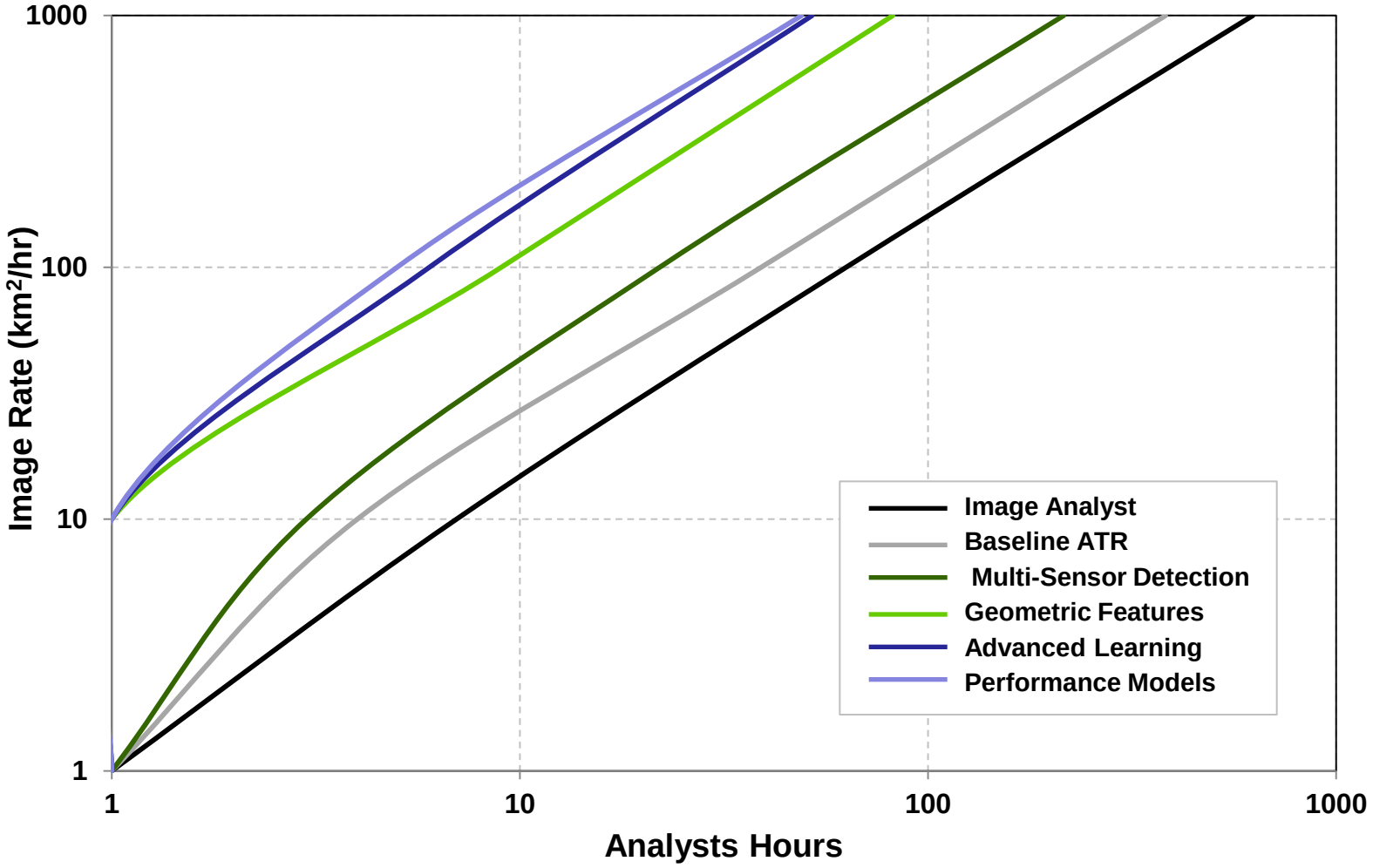
Tracking Analysis



Imagery Processing Chain



Detection/Classification Analysis



Text Analysis

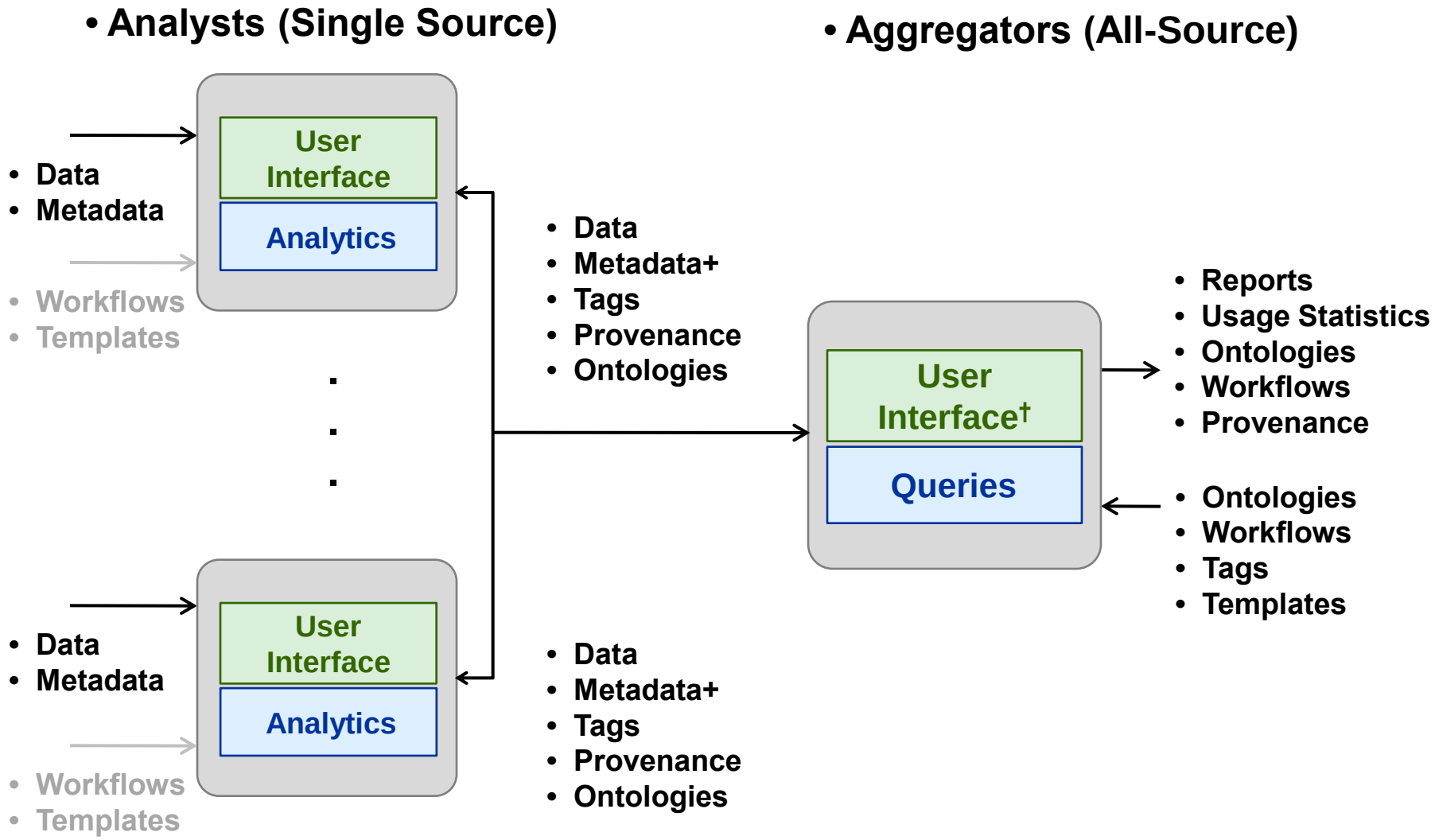
	Advanced Machine Translation	HSCB Analysis	Document Exploitation (DOCEX)*	A&V from Text (Link/Temporal /Spatial)
Text Preparation (OCR, Speech, MT, Zoning)				
Entity/Event Resolution & Consolidation				
Advanced Entity/Relation/Event Extraction				
Time/Location Stamping				
Subjectivity/Sentiment Extraction				
Text Mining				
Portability (Genre/Domain/Language)				
Multilingual Extraction				

* This refers *to operational* Document Exploitation *(*DOCEX); when Special Ops Forces (SOF) finds hard copy documents at a site and we need to process for intel info

Acronyms & Abbreviations
A&V = Analysis & Visualization
HSCB = Human Social Cultural Behavioral
MT = Machine Translation
OCR = Optical Character Recognition

N/A TRL 6-9 TRL 3-6 TRL 1-3

User Interface Layer



†End User Programmable

User Interaction Layer

- **Problem Statement: Existing interface tools do not support the user's need to collaborate, visualize, adapt and manage knowledge gained from sensing assets**

- **3-to-5 year timeframe objective**
  - User tools that aid data discovery, link communities, support aggregation and provide natural user interfaces

- **7-to-10 year timeframe objective**
  - Never-ending learning systems that maintain and reason over millions of facts to identify new knowledge
  - Workflow tools that capture and teach analysts' best practices

Summary

- The Data-to-Decisions program develops technology for the rapid development of flexible new Decision Support Systems
- Program consists of a series of relevant challenge problems that advance the underlying technology in data management, analytics and user interfaces
- Execution is through a consortium that addresses the challenge problems in a coherent and integrated team approach
- Major research initiatives focus on developing extendable analytic approaches and advanced user-interface modules